REMARKS

Claims 1-22 are pending in the present application. Claims 1, 4, 9-11, 14-16, 19, 20, 22 are amended and claims 2-3. 5-8, 12-13, and 17-18 are cancelled without prejudice. Reconsideration of the claim rejections are respectfully requested in view of the following remarks.

Claim Rejections - § 101

Claims 15-22 are rejected under 35 U.S.C. 101. The Examiner contends that claims 15-22 are directed to non-statutory subject matter because they are directed to automatic speech recognition systems and the specification states that such systems can only be implemented in software.

Applicants respectfully disagree. The specification never restricts automatic recognition systems to only software implementations. Page 5, lines 18-19 of the specification states that "the system and method described herein may be implemented in various forms of hardware". Further, while, page 5, lines 21-22 of the specification states that "[p]referably, the present invention is implemented in software", a preference for a software implementation does not exclude hardware implementations.

Withdrawal of the rejections under 35 U.S.C. 101 is respectfully requested.

Claim Rejections - § 103

Claims 1-22 stand rejected under 35 U.S.C 103(a) as being unpatentable over U.S. Patent No. 5,995,930 to Hab-Umbach in view of U.S. Patent No. 6,078,885 to Beutnagel, as set forth in pages 2-4 of the Office Action.

Claims 1-8

Beutnagel uses text-to-speech (TTS) to communicate a potential pronunciation to a human and then the human accepts it or hand-edits the written pronunciation and re-listens to see if the new version is better (see col. 4, lines 12-27). However, embodiments of the present invention automatically analyze the output of the TTS to help improve the recognition system, i.e., there is **no human** in the loop.

Claim 1 has been amended to better clarify how one embodiment of this automatic analysis is performed. For example, claim 1 has been amended to recite *for each synthetic waveform, time-aligning feature vectors of the synthetic waveform with feature vectors of the original waveform at a phoneme level, computing a mean of the feature vectors which align to each phoneme for the original waveform and the synthetic waveform, computing a distance measure between each phoneme mean of the original waveform and the synthetic waveform, summing the distance measures to generate an overall distance measure representing a distance between the original waveform and the synthetic waveform, and selecting for output the textual transcription corresponding to the synthetic waveform having a smallest overall distance measure.*

Beutnagel does **not** teach the human or any other mechanism performing for each pronunciation: **i)** time-alignment of feature vectors of the pronunciation with feature vectors of an original waveform at a phoneme level, **ii)** computing a mean of the feature

vectors which align to each phoneme for the original waveform and the pronunciation, **iii)** computing a distance measure between each phoneme mean of the original waveform and the pronunciation, **iv)** summing the distance measures to generate an overall distance measure representing a distance between the original waveform and the pronunciation, and **v)** selecting for output a textual transcription corresponding to the pronunciation having a smallest overall distance measure.

Further, the deficiencies of Beutnagel in this regard are not cured by Hab-Umbach. Hab-Umbach teaches (in col. 5, lines 27-30) comparison of test signals with reference signals. However, Hab-Umbach does **not** teach its comparison being performed **i)** by time-aligning feature vectors of a test signal with feature vectors of a reference signal at a phoneme level, **ii)** by computing a mean of the feature vectors which align to each phoneme for the test signal and the reference signal, **iii)** by computing a distance measure between each phoneme mean of the test signal and the reference signal and **iv)** by summing the distance measures to generate an overall distance measure representing a distance between the test signal and the reference signal. Further, Hab-Umbach does not teach selection of a textual transcription corresponding to the reference signal having a smallest overall distance measure.

For at least the foregoing reasons, the combination of Beutnagel and Hab-Umbach fails to disclose or suggest claim 1; and thus, claim 1 is believed to be patentable over said combination.

Claims 3-8 are cancelled without prejudice.

Claim 2 is believed to be patentable over said combination at least by virtue of its dependence from claim 1.

Claims 9-14

Claim 9 has been amended to better clarify how another embodiment of the above described automatic analysis is performed. For example, claim 9 has been amended to recite *for each synthetic waveform, computing a distance measure between the synthetic waveform and the original waveform, summing the distance measures to generate an overall distance measure representing a distance between the original waveform and the synthetic waveform, generating a score from the overall distance measure, an acoustic model score for the synthetic wave, and a language model score of the synthetic waveform*, as recited in amended claim 9.

As discussed above, Beutnagel uses a human listener to select a best pronunciation. Further, there is no teaching in Beutnagel of the human or any other mechanism generating a score based on an overall distance measure between a pronunciation and an original signal, an acoustic model score of a synthetic wave, and a language score of a synthetic wave. While Hab-Umbach teaches (in col. 5, lines 25-30) comparison between a test signal and a reference signal to output scores, there is no teaching in Hab-Umbach of these scores being based on a based on an overall distance measure between a pronunciation and an original signal, an acoustic model score of a synthetic wave, and a language score of a synthetic wave.

For at least the foregoing reasons, the combination of Beutnagel and Hab-Umbach fails to disclose or suggest claim 9, and thus claim 9 is believed to be patentable over said combination.

Claims 12-13 are cancelled without prejudice.

Claims 10-11 and 14 are believed to be patentable over said combination at least by virtue of its dependence from claim 9.

Claims 15-22

Claim 15 has been amended to better clarify how another embodiment of the above described automatic analysis is performed. For example, claim 15 has been amended to recite *a means to perform a speaker normalization on the original waveform to match vocal-tract characteristics of a speaker from whose data the TTS is derived, computing a distance measure between the synthetic waveform and the normalized original waveform, and summing the distance measures to generate an overall distance measure representing a distance between the normalized original waveform and the synthetic waveform.*

As discussed above, Beutnagel uses a human listener to select a best pronunciation. There is no teaching in Beutnagel of the human or another mechanism performing normalization on each pronunciation to match vocal-tract characteristics of the human listener. As discussed, Hab-Umbach teaches comparison of a reference signal to a test signal. However, there is no teaching in Hab-Umbach of performing any such speaker normalization on either the reference signal or the test signal.

For at least the foregoing reasons, the combination of Beutnagel and Hab-Umbach fails to disclose or suggest claim 15, and thus claim 15 is believed to be patentable over said combination.

Claims 17-18 are cancelled without prejudice.

Claims 16 and 19-22 are believed to be patentable over said combination at least by virtue of its dependence from claim 9.

Withdrawal of the rejections under 35 U.S.C. 103(a) is respectfully requested.

Conclusion

In view of the foregoing amendments and remarks, it is respectfully submitted that all the claims now pending in the application are in condition for allowance. Early and favorable reconsideration is respectfully requested.

Respectfully submitted,

Dated: 4/21/09 By: Robert J Newman

Robert J Newman
Reg. No. 60,718
Attorney for Applicants

F. CHAU & ASSOCIATES, LLC
130 Woodbury Road
Woodbury, NY 11797
Telephone: (516) 692-8888
Facsimile: (516) 692-8889